

# Genomics of *Ciona intestinalis*

- David N. Keys
  - dnkeys@lbl.gov
- Joint Genome Institute
  - DOE • LBNL • LLNL
  - Walnut Creek, CA
  - [www.jgi.doe.gov](http://www.jgi.doe.gov)



## A Functional Genomics Approach to Developmental Genetics

**Joint Genome Institute**  
**LBNL - LLNL - DOE**

**University of California**  
**Berkeley**

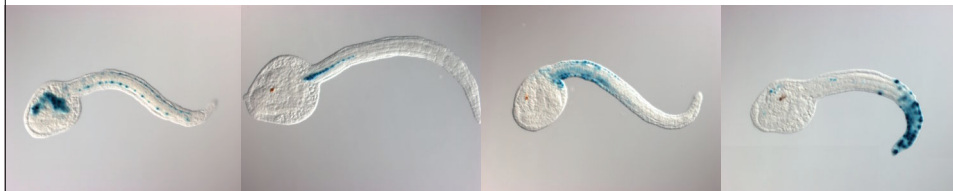
Byung-in Lee  
Chris Detter  
Stephan Trong

Syvia Ahn  
Dave Engle  
Orsalem Kahsai

Naoe Harafuji  
Anna Di Gregorio  
Mike Levine

Dan Rokhsar  
Paul Richardson  
Trevor Hawkins

Maria Shin  
Joann Wang  
Mei Wang



## *Ciona intestinalis* as an experimental organism

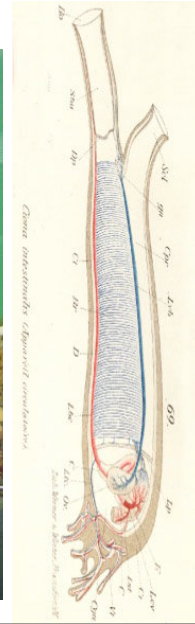


Basal Chordate  
Genomic simplicity  
Easy transgenics  
Ectopic expression  
*Cis*-regulation studies

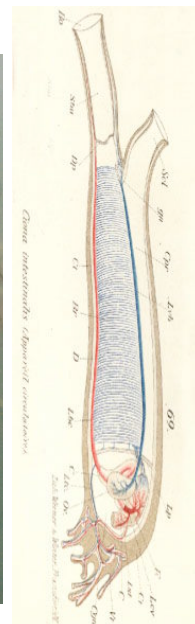
## *Ciona intestinalis* as an experimental organism



Basal Chordate  
Genomic simplicity  
Easy transgenics  
Ectopic expression  
*Cis*-regulation studies

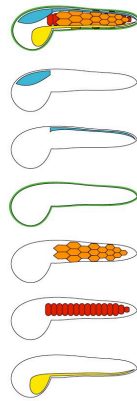


## A close-up photograph showing a person's hand and forearm reaching into shallow, rippling water. The person is wearing a black digital sports watch with a black strap. The hand is positioned near the water's surface, with fingers slightly curled. The water is clear, revealing a rocky riverbed with various sized stones and pebbles in shades of brown, tan, and grey. The water's surface is covered with small, concentric ripples, indicating a gentle current or the movement of the hand. The background is slightly blurred, focusing attention on the hand and the immediate water surface.



# *Ciona intestinalis*

## Larval body plan



*Ciona intestinalis*

Cerebral Vesicle/Brain

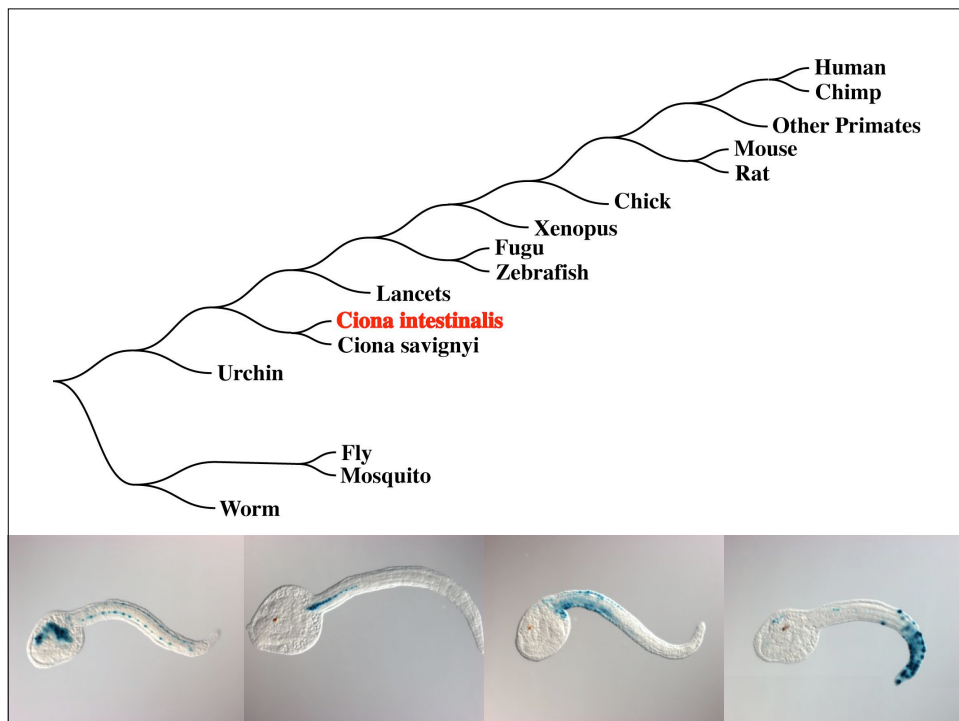
Neural Tube

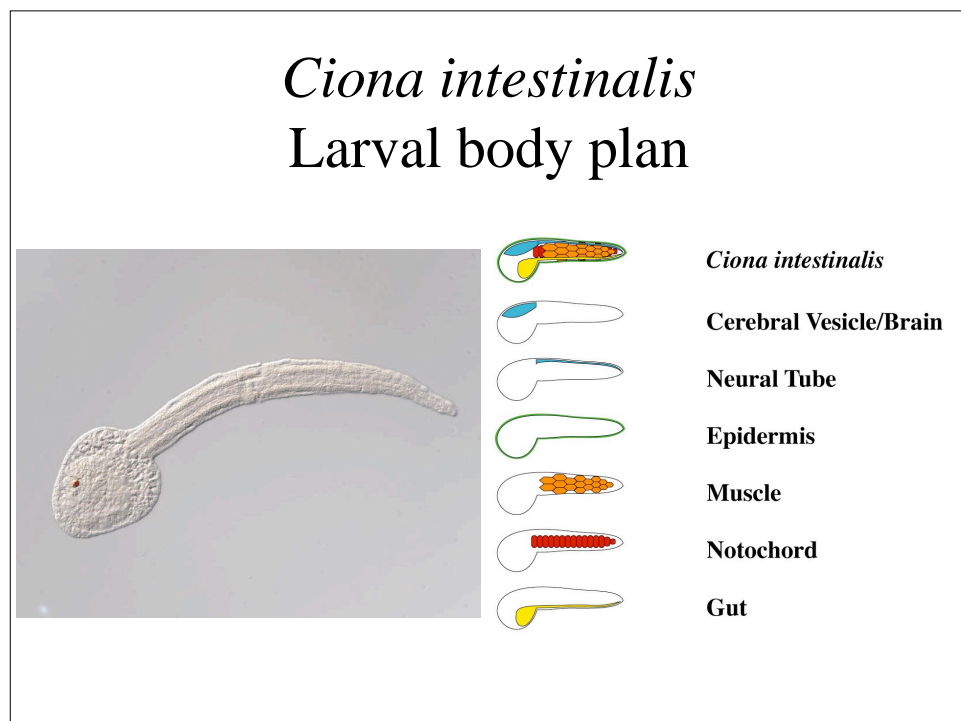
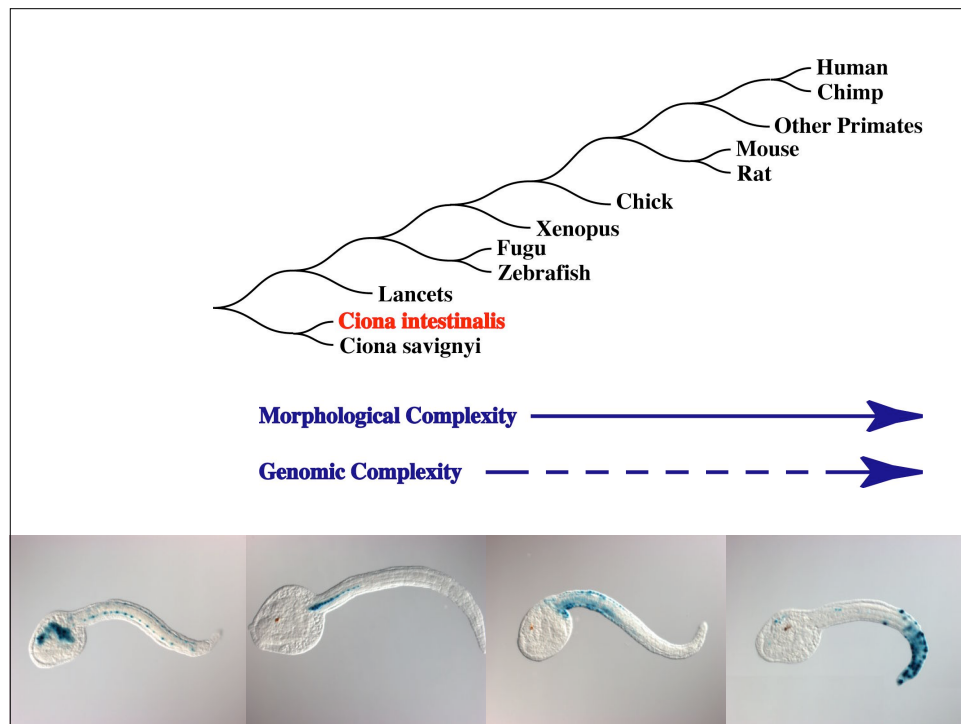
Epidermis

Muscle

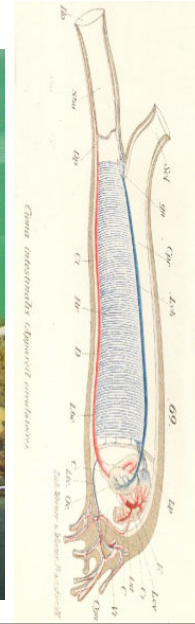
Notochord

Gut

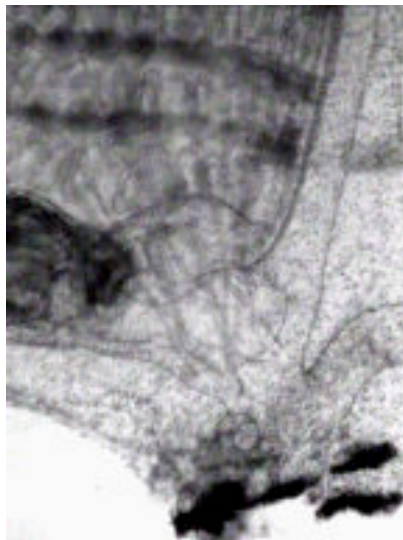




## *Ciona intestinalis* Adult body plan



## Adult Circulatory System



- Heart
  - Simple single chamber
  - Muscle & pericardium
  - Asymmetrical looping
  - Reversible
- Vascular System
  - Two major vessels
  - Open sinuses
- Blood
  - Eight known cells types



## Genomic Simplicity



- ~160,000,000bp
- ~16,000 genes
- ~1 gene every 10kb
- ~500,000 ESTs
- Small gene families
  - Predates vertebrate duplication
  - Little redundancy
  - Easily identified orthologs

## Experimental Tractability



### **Disadvantage**

No true genetics

### **Advantages**

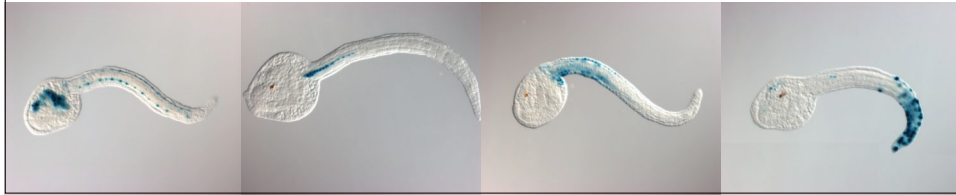
Easy transgenics

Scorable phenotypes

Availability

## Functional Genomics

- Studying large sets of genes in parallel rather than single genes
- Experimental, not observational or modeled
- Invent new hypothesis testing experiments
- Scale traditional hypothesis testing experiments to the entire genome



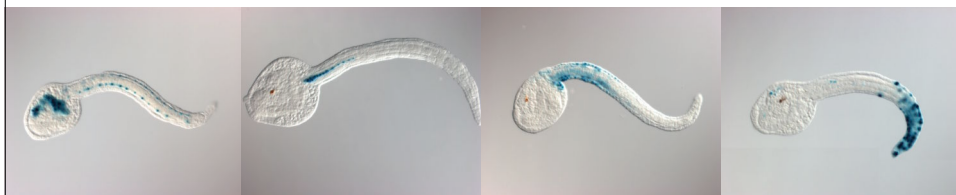
## Large Scale *Cis*-Reg Hunts

### •Primary Goal

- Screen genomic libraries for *cis*-regulatory activity
- Catalog a large number of functionally defined *cis*-regulatory elements

### •Secondary Goal

- Do some targeted developmental genetics along the way

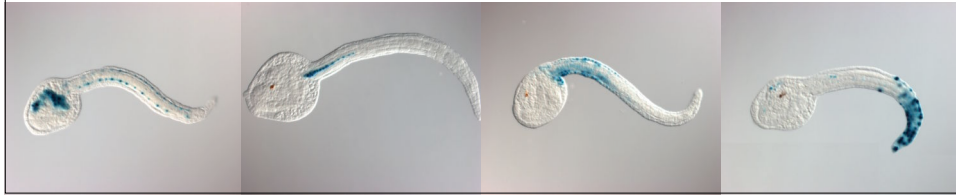




## Results

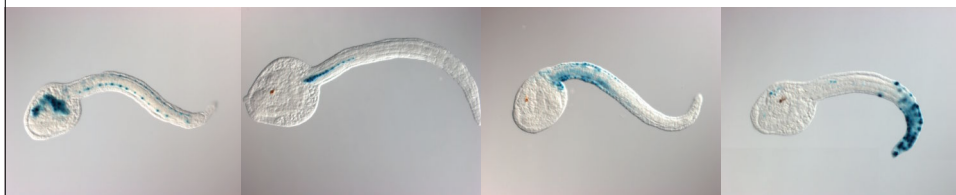
### Catalog a large number of functionally defined *cis*-regulatory elements

- Design, implementation and results of a small scale pilot screen of random genomic DNA - 11
- Design, implementation and results of an exhaustive screen of a medium size (250kb) genomic domain
- Design, implementation of an on going large scale screen of random genomic DNA



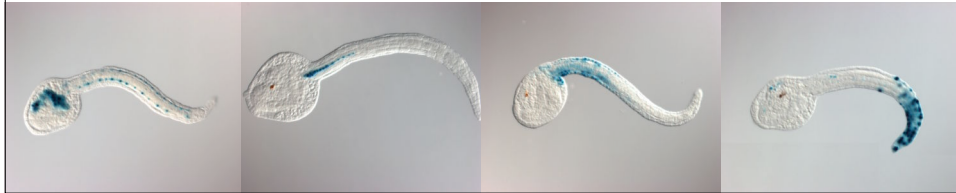
## Themes to Keep in Mind

- Trade Offs
  - Number of characterized elements
  - Resolution of the characterizations
- Biases
  - Experimental biases
  - Experimenter's biases
    - Nature of enhancers vs detection methods

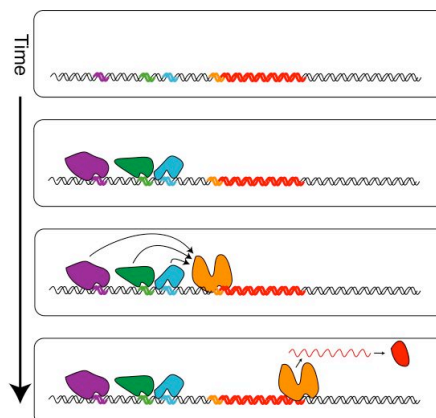


# Technology to take *cis*-regulatory screening to the genomic level

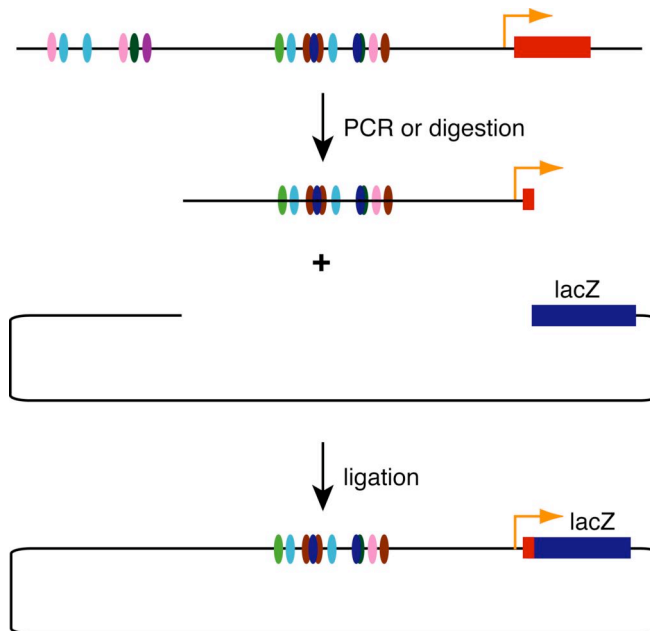
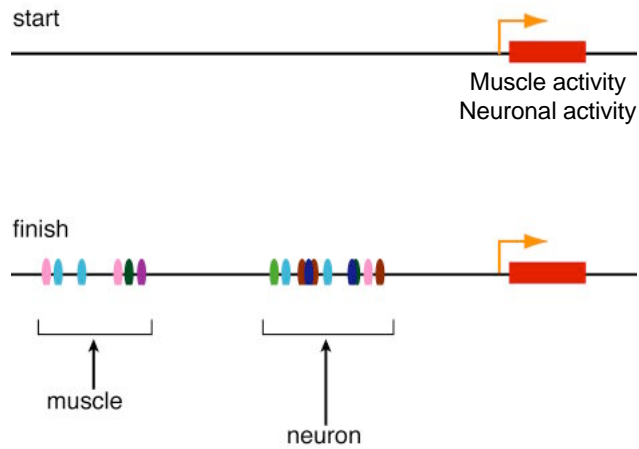
Scale traditional hypothesis testing experiments to the entire genome

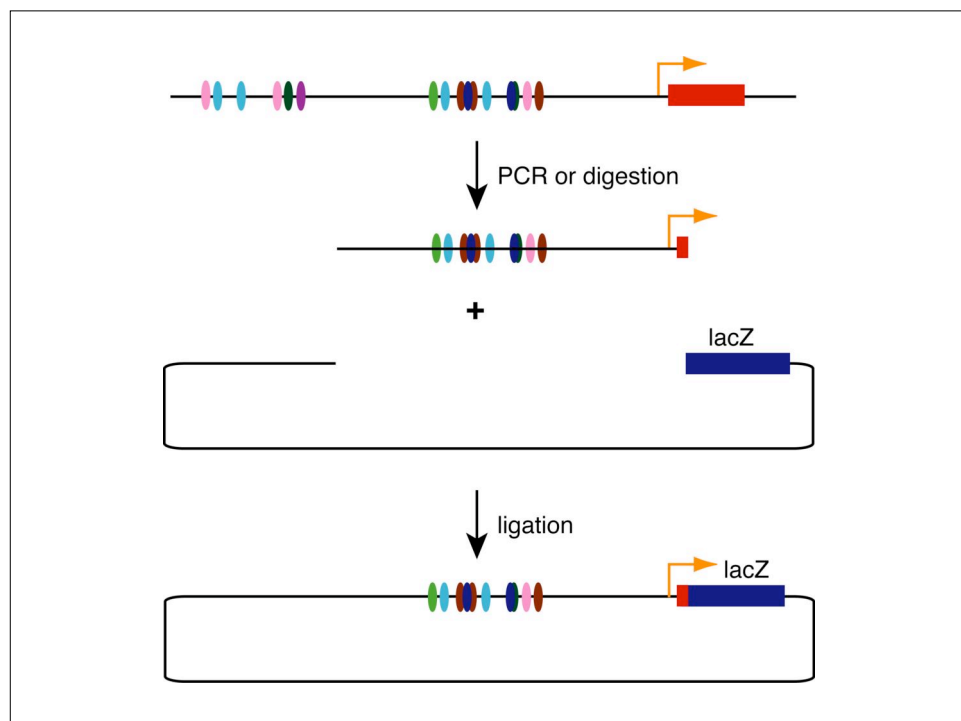
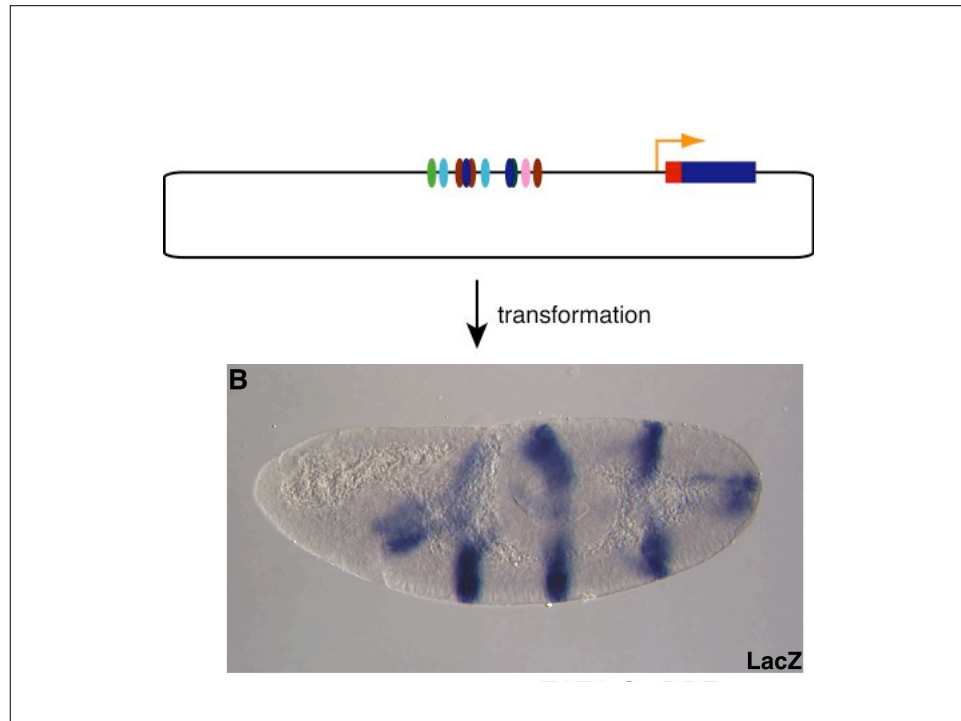


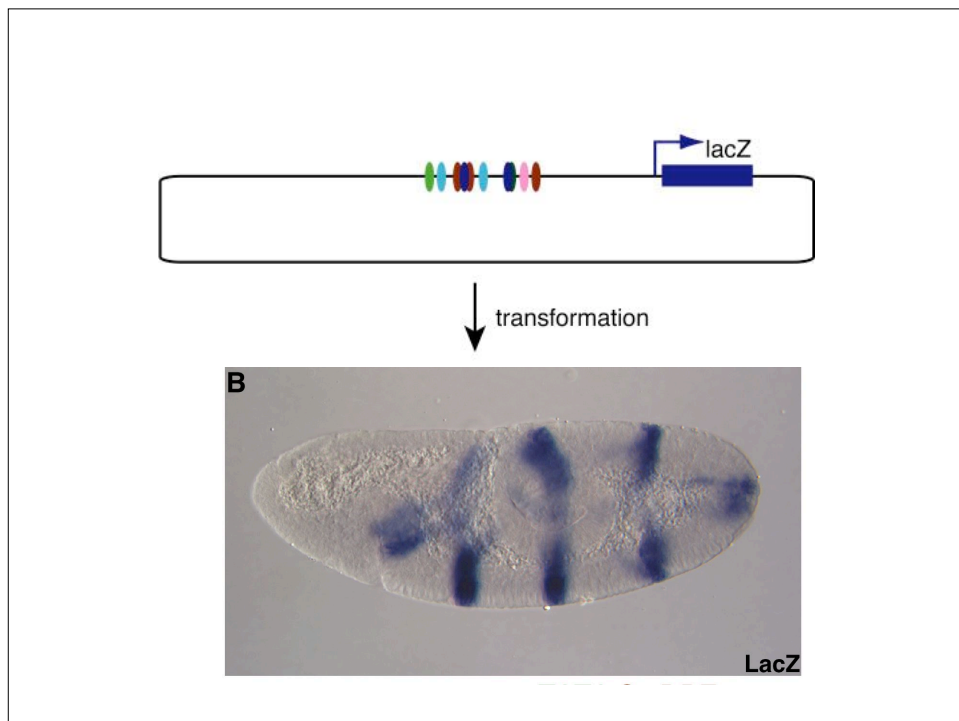
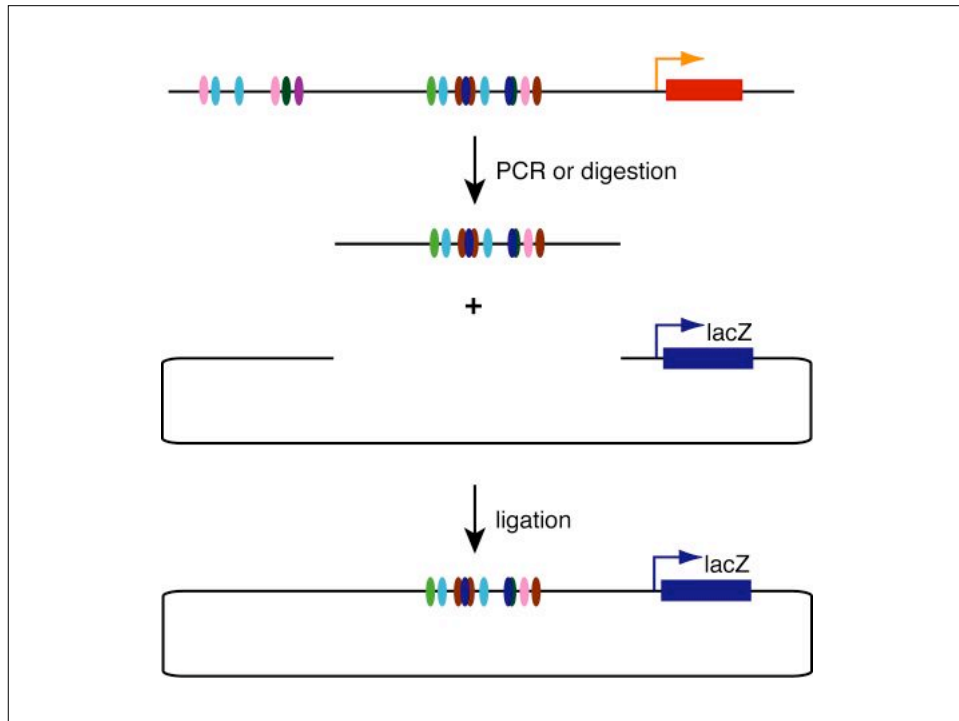
## A genetic switch



# Enhancer Characterization







## *Drosophila* transformation

- Collect naturally laid eggs
- Dechoriation
- Transform by single embryo microinjection
- Individually rear to 2nd generation
- Screen
- **Total Time: month(s)**

## Traditional Enhancer Characterization

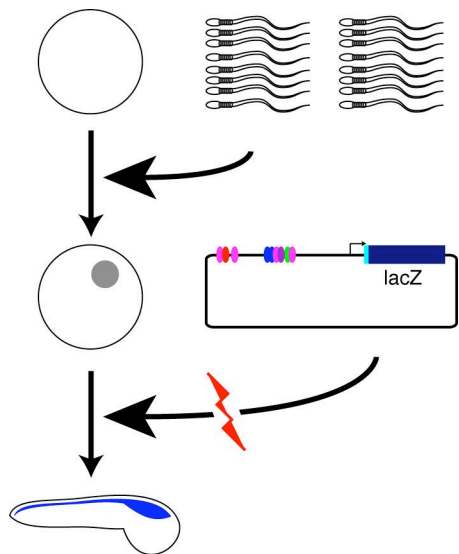
- Targeted
- Slow/Labor intensive
  - Building specific DNA constructs
  - Transforming into animals
  - Maintaining/screening animals



## *Ciona intestinalis* Sea Squirts



### *Ciona intestinalis* transformation



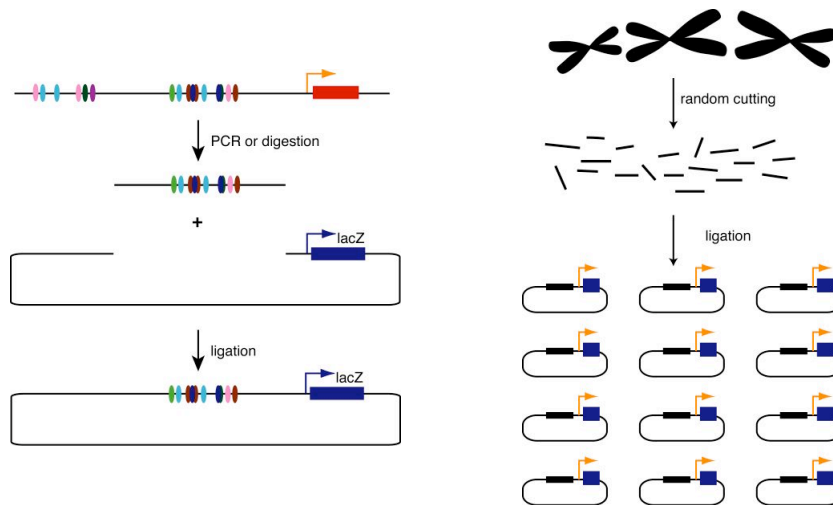
- Mix eggs & sperm
- Add a single plasmid
- Transform by electroporation
- Incubate 8-24 hours
- Stain (LacZ/GFP/in situ)
- Visually screen

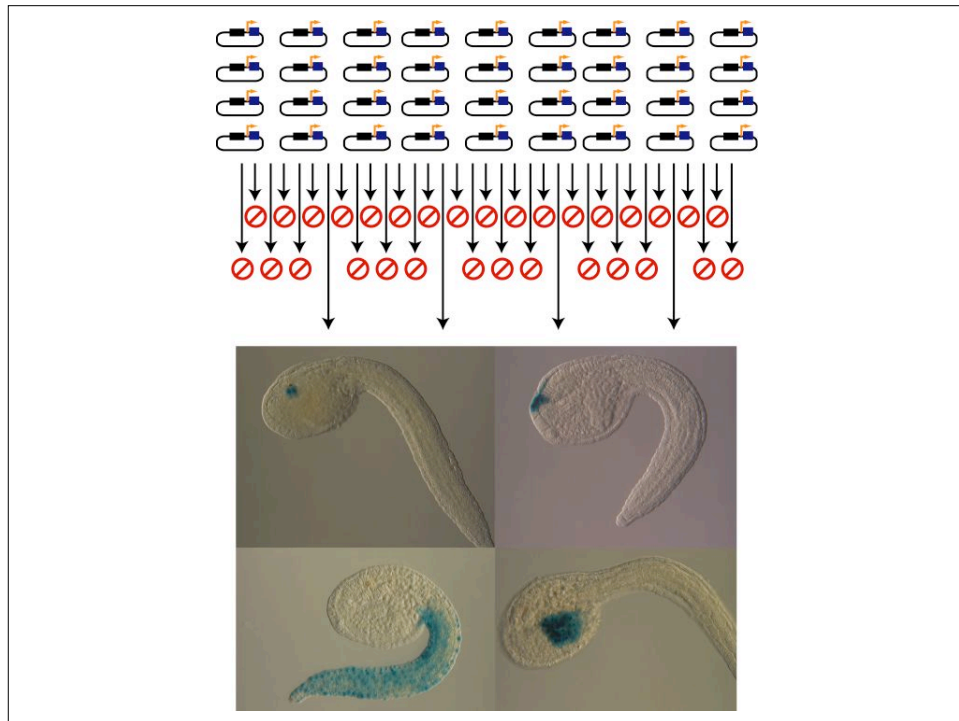
- **Total time: 24 hours**

## Standard *Ciona* Enhancer Characterization

- Targeted
- Slow/Labor intensive
  - Building specific DNA constructs
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals

## Build enhancer screening libraries instead of specific constructs



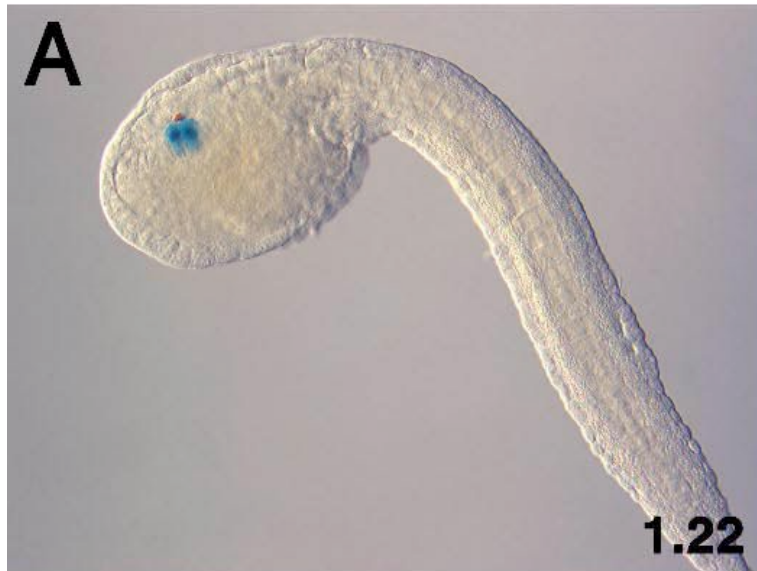


## *Ciona* Enhancer Screening

- Non-Targeted
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals
  - Building random DNA constructs
- Limiting factors
  - DNA preps (50-100ug)
  - Transformation window (single cell embryos)
  - Imaging

## Pilot Genomic Screen

- Construct:
  - *Ciona Forkhead* basal promoter
  - *lacZ* marker detected by *beta*-Gal activity
  - Random genomic Sau3AI frags, 1.7kb average
- Prediction:
  - Will find cis-regulatory DNA
  - Gene density = 1 gene per 10kb. Therefore could find 1 enhancer every 10kb





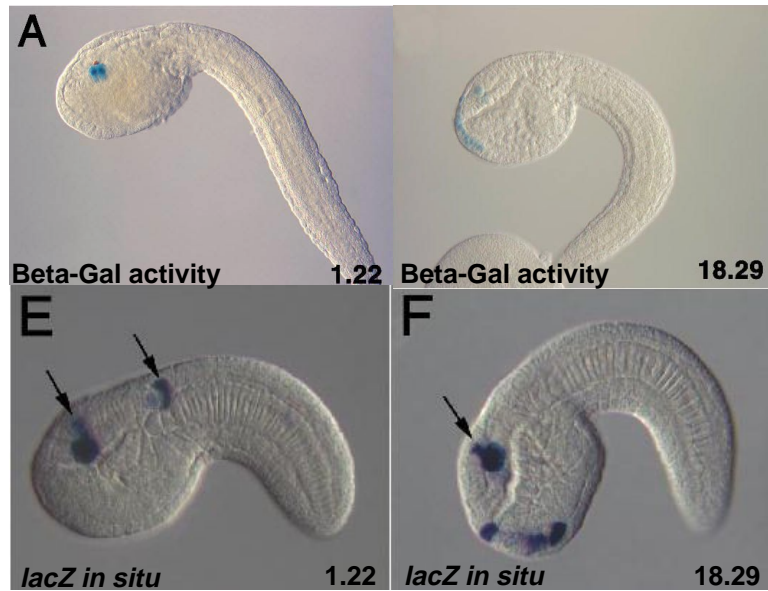


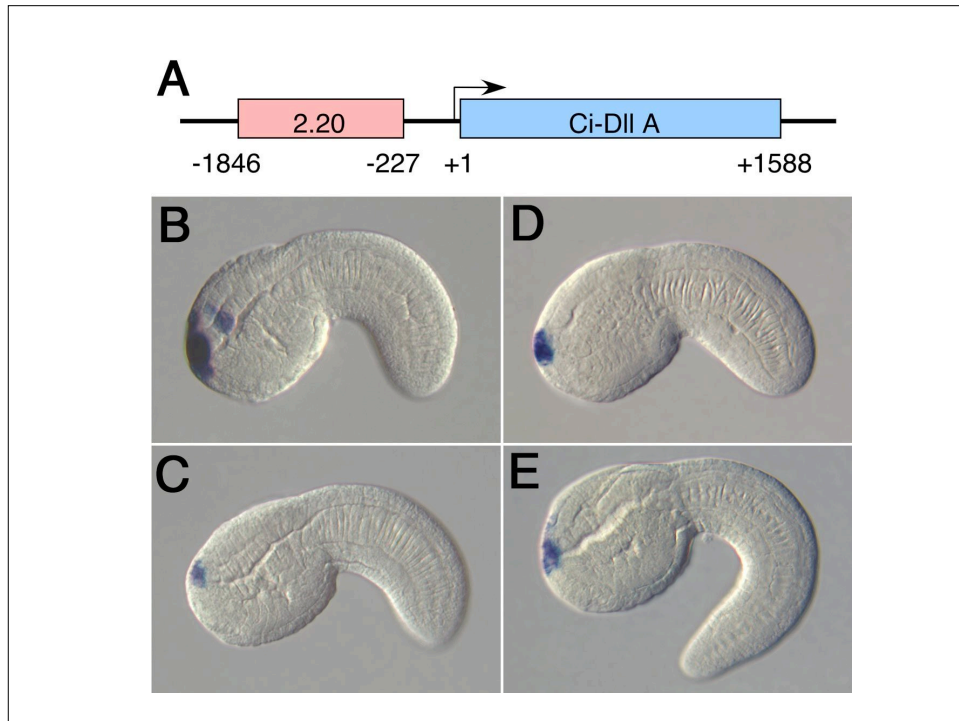




## First genomic screen

- 138 constructs
- 250kb screened
- 0.15% of the genome
- **Results:**
  - 11 strong cis-regulatory elements
  - At least 8 appear to be “*real*” enhancer elements
  - One confirmed enhancer
  - 1 detectable element every 23-31 kb
  - 1 detectable every 2-3 genes.





- **Prediction:**

- Gene density = 1 gene per 10kb
- Therefore could find 1 enhancer every 10kb

- **Results:**

- 1 detectable element every 23-31 kb
- 1 detectable element every 2-3 genes

## Potential Issues

- Promoter specificity
- Insulators & repressors
- Enhancer Polarity
- Promoter competition
- Enhancers fragmented during cloning
  
- Timing
- Insufficient detection strength

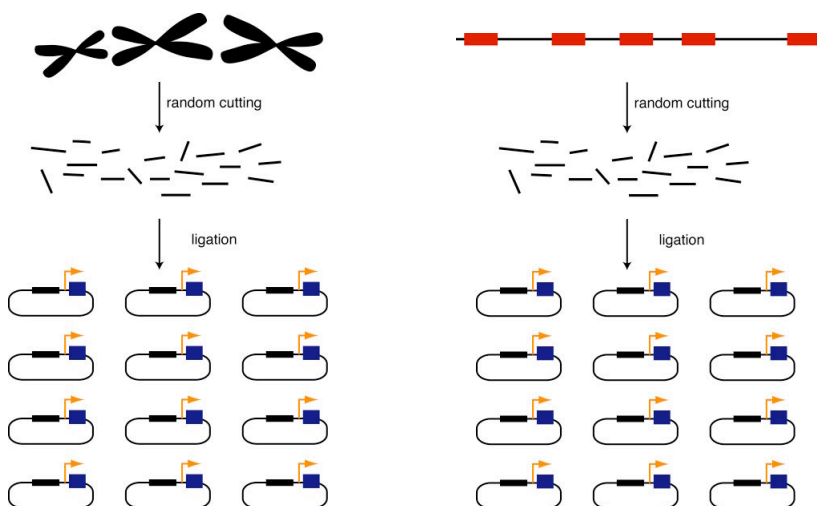
## *Ciona* Enhancer Screening

- Non-Targeted
- Fast/Not labor intensive
  - Transforming into animals
  - Screening animals
  - Building random DNA constructs
- Limiting factors
  - DNA preps (50-100ug plasmid)
  - Transformation window (single cell embryos)
  - Imaging

## Limiting Factors

- DNA preps (50-100ug plasmid)
  - Qiagen Midipreps - up to 48 constructs per day
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
- Imaging
  - Quality trade offs - Tough decisions

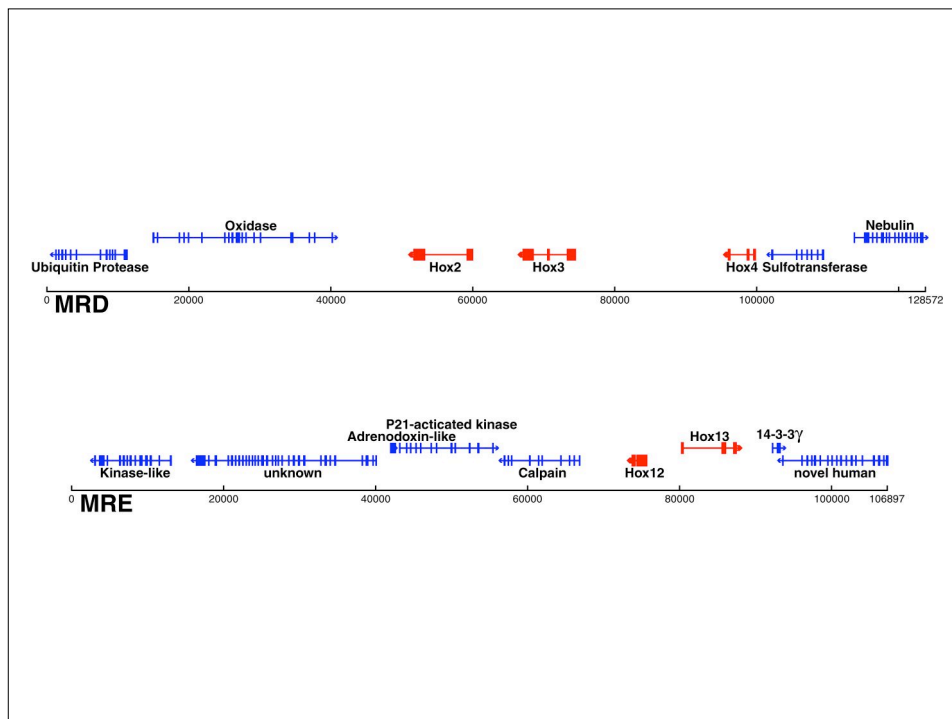
## Semi-targeted *Ciona* enhancer screen Build random libraries from limited regions



## Target: *Ciona Hox Complex*

### Predictions:

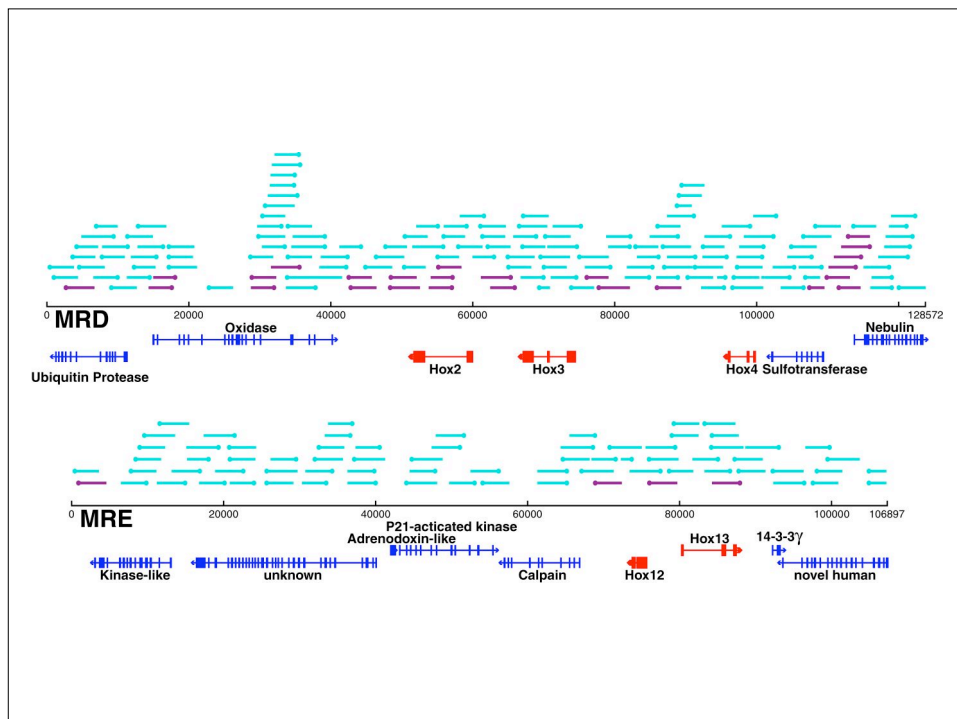
- Should be a single *Hox Complex*
- Should be a single domain
- Predictable expression patterns
  - *Hox3* & *Hox5* described by in situ  
(Branno & Di Lauro, Stazione Zoologica, Naples)

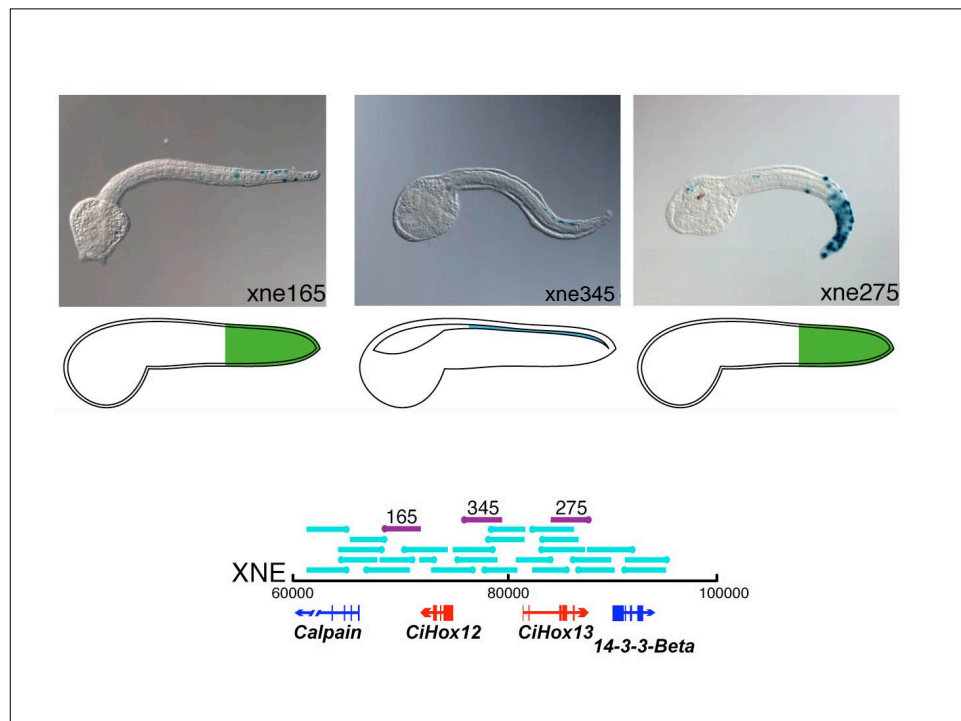
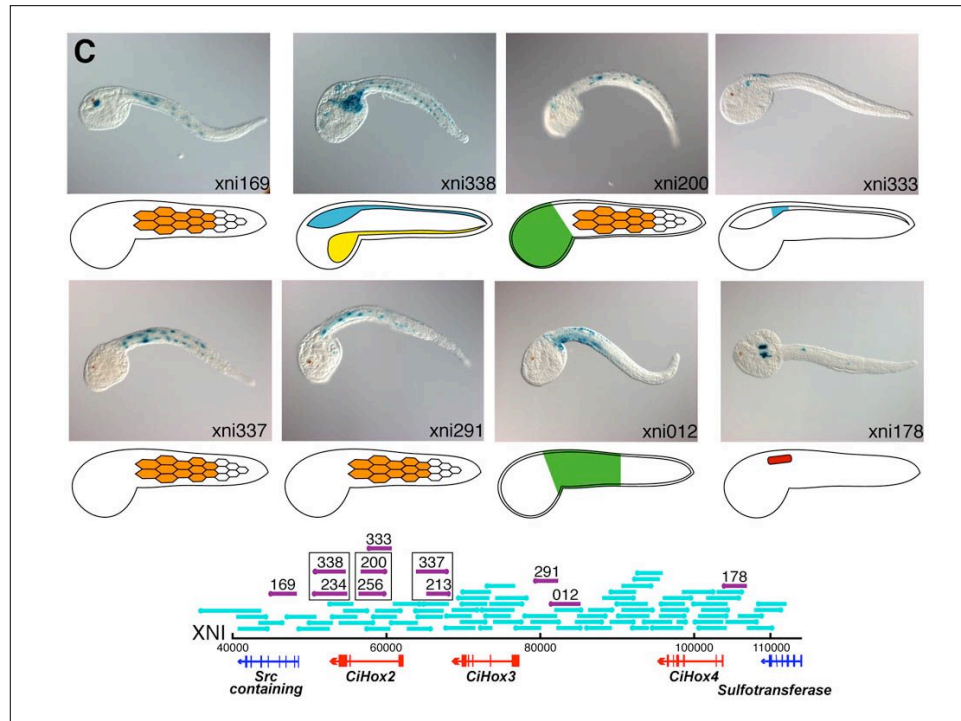


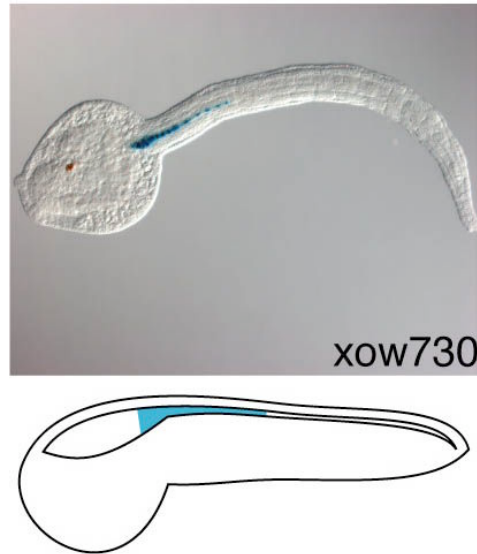


## *Ciona Hox Complex*

- Should be a single *Hox Complex*
  - Correct
- Should be a single domain
  - Wrong, at least 4 separable domains
- Predictable expression patterns



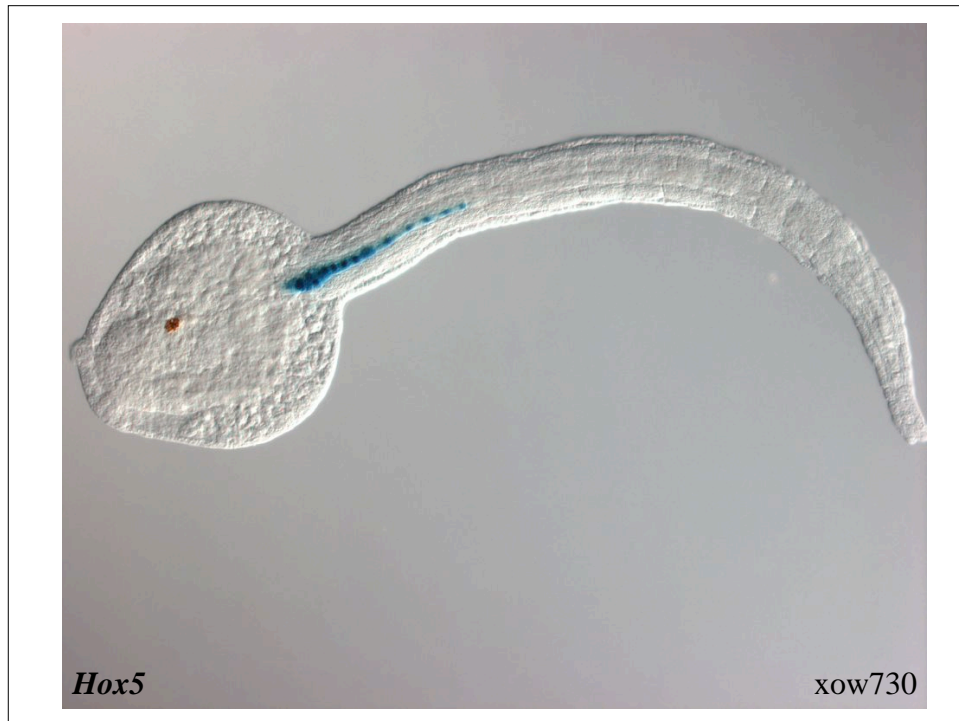




## Primary results

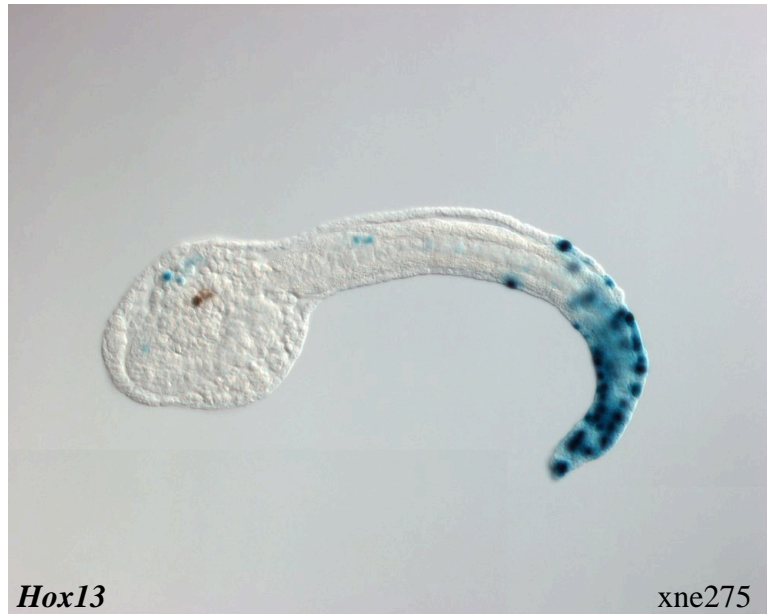
- 221 clones electroporated & passed
- 39 clones showed positive signal
- Range for enhancers actually found:
  - Likely Maximum 30
  - Likely real 21
  - Minimum 17
  - Likely *Hox* 08











## *Ciona Hox* Complex

- Should be a single *Hox* Complex
  - Correct
- Should be a single domain
  - Wrong, At least 4 separable domains
- Predictable expression patterns
  - Correct, Nested CNS
  - Unexpected, Nested Epidermis

## Limiting Factors

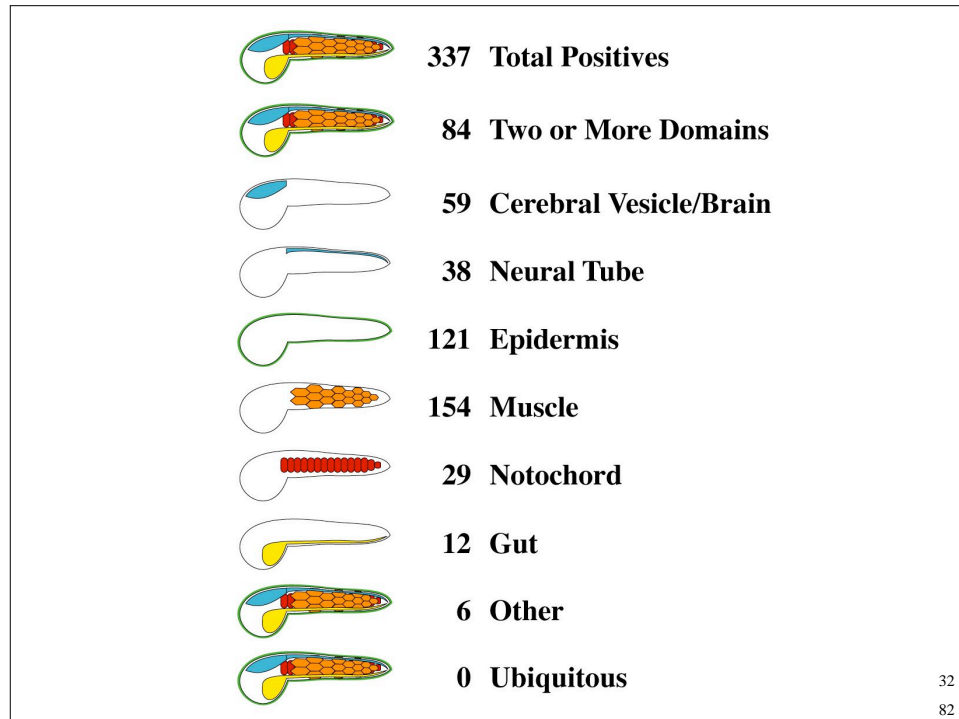
- DNA preps (50-100ug plasmid)
  - Qiagen Midipreps - up to 48 plasmids per day
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
- Imaging
  - Quality trade offs - Tough decisions

## Full Genome Scale Up



## Limiting Factors

- DNA preps (50-100ug plasmid)
  - Rolling Circle Amplification
- Transformation window (single cell embryos)
  - 24 separate constructs per batch
  - 480 constructs per week
- Imaging
  - Quality trade offs - Tough decisions
  - Automation??

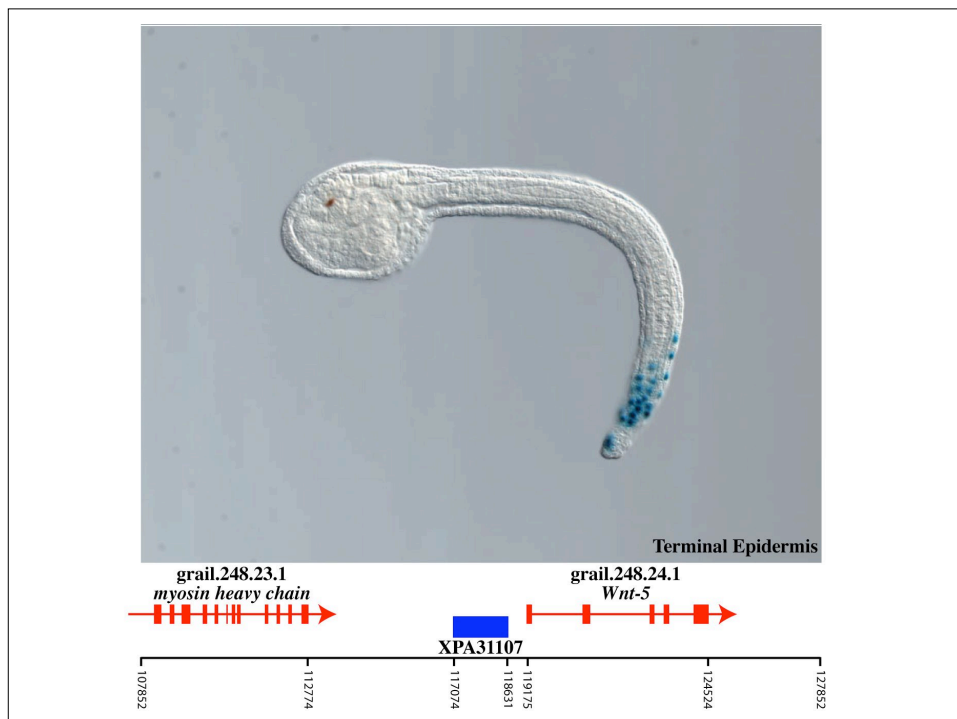


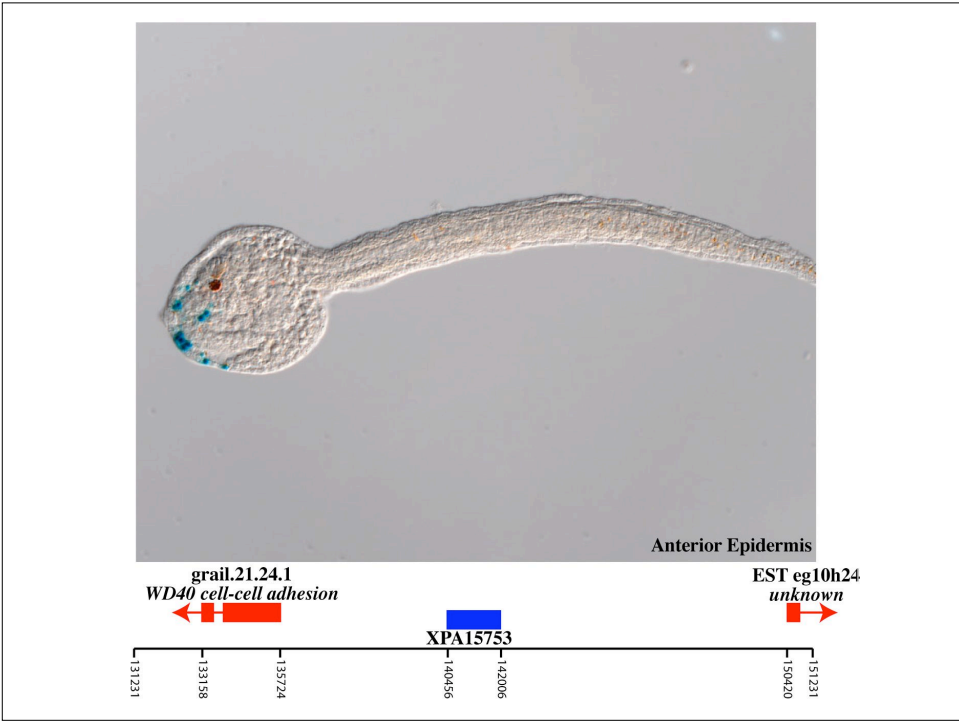
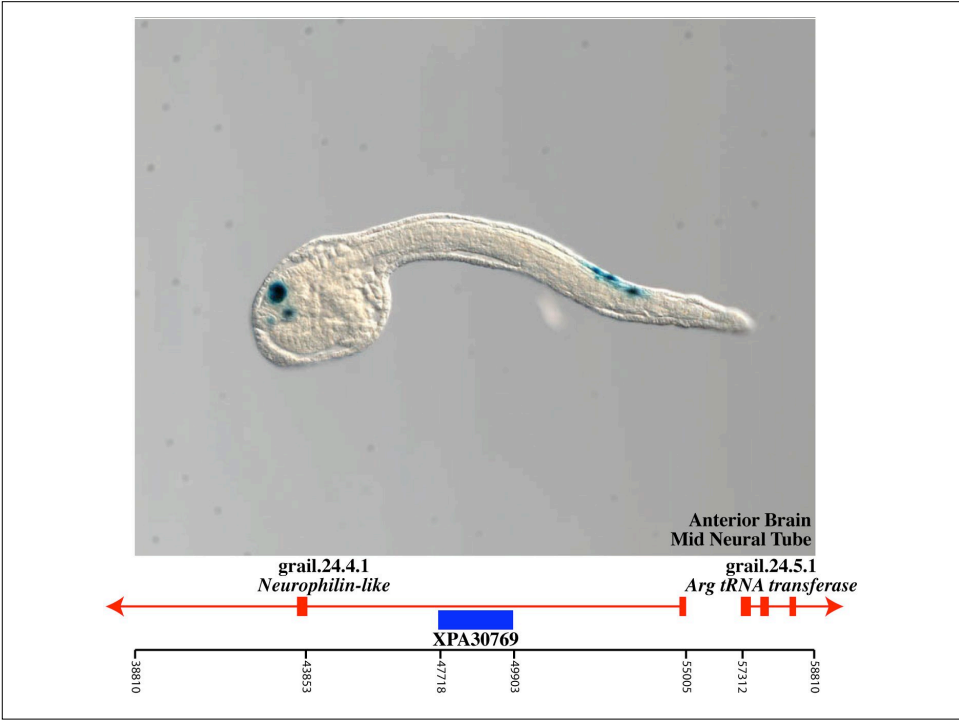
## Scale Up

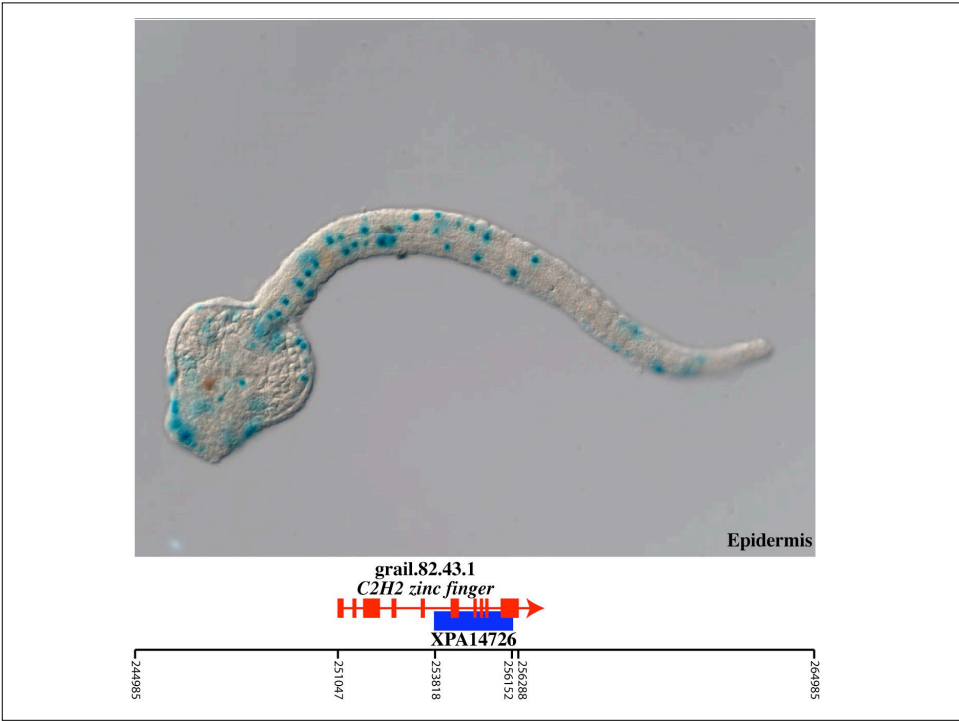
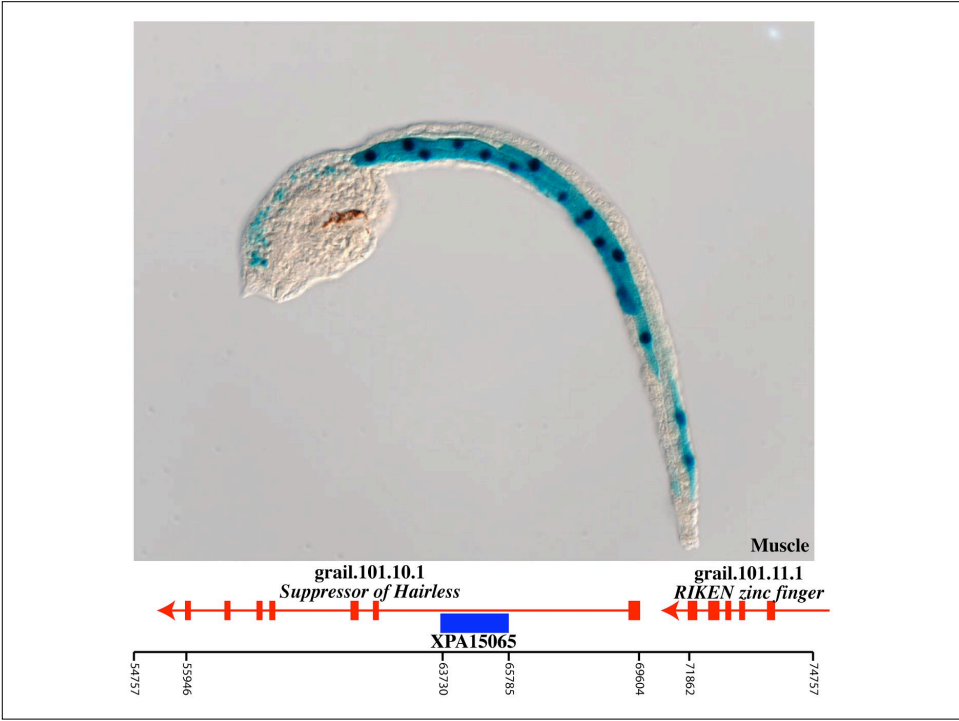
- XPA28186 All epidermis Hypotheical 109.7 kDa protein
- XPA28213 Tail Muscle Serine/Threonine Kinase MASK
- XPA28241 Ventral Mid Brain Homolog to cDNA FLJ10540
- XPA28134 Notochord Low Sequence quality
- XPA30404 Tail Muscle RAR Related Steroid Receptor
- XPA30769 Dorsal Brain, Neural Tube Arginine tRNA protein transferase
- XPA30770 Muscle & Notochord Proline Oxidase 1
- XPA31107 Post Tail Epidermis Wnt-2
- XPA28831 All CSN & Epidermis MORN motif containing
- XPA28492 Single Cell in Brain unknown but conserved protein
- XPA28855 Post brain & Neural Tube Protein kinase Ck2-beta
- XPA29631 Neural Tube, All Gut unknown but conserved protein
- XPA25239 Unknown cells in head MEC-8 like

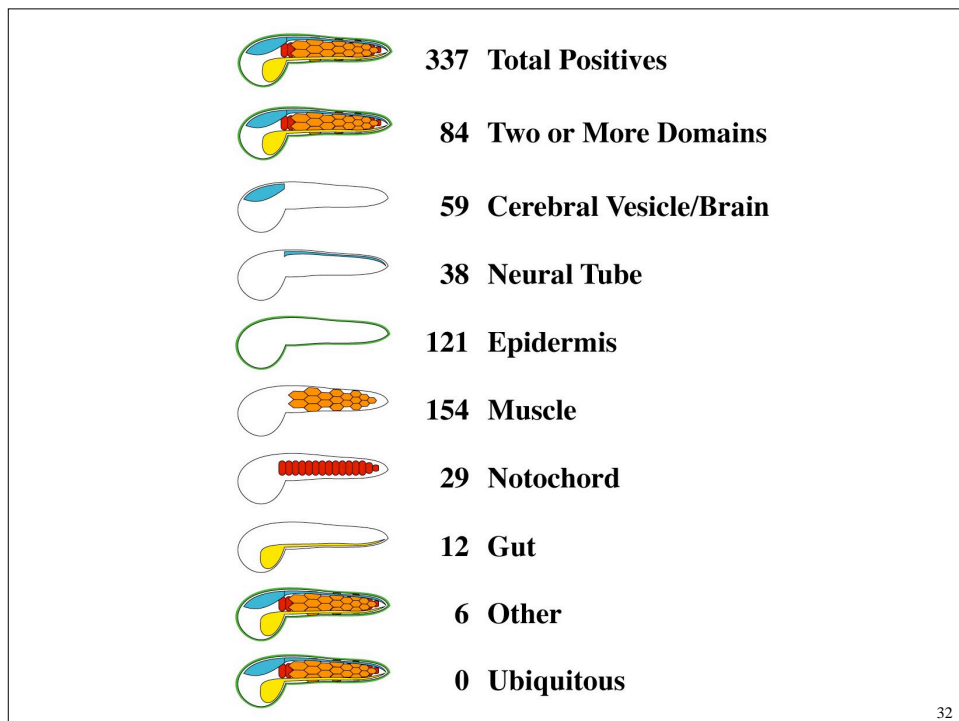
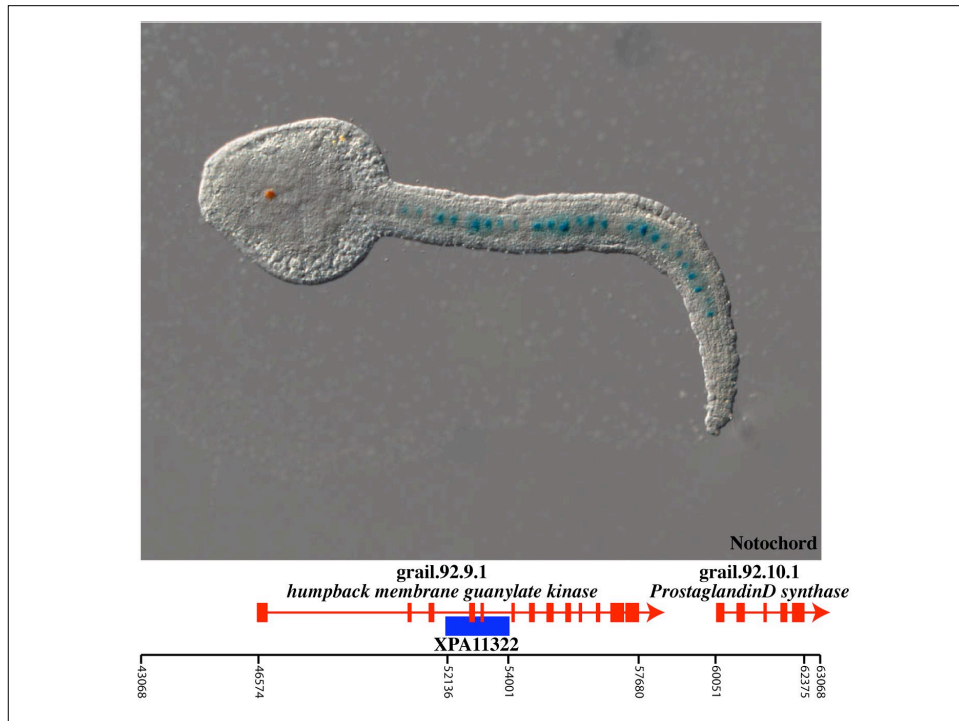
# Genomic Integration

- For most random constructs, 2 end runs will
  - Identify entire subcloned sequence
  - Identify both flanking ORFs
  - Tie into EST in situ project









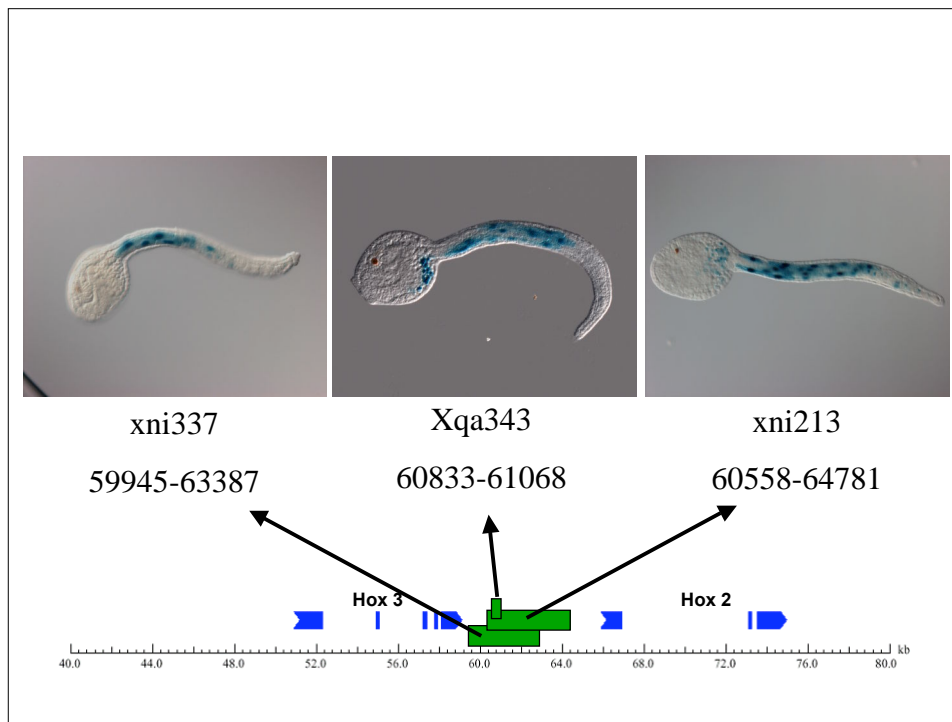


What's next?

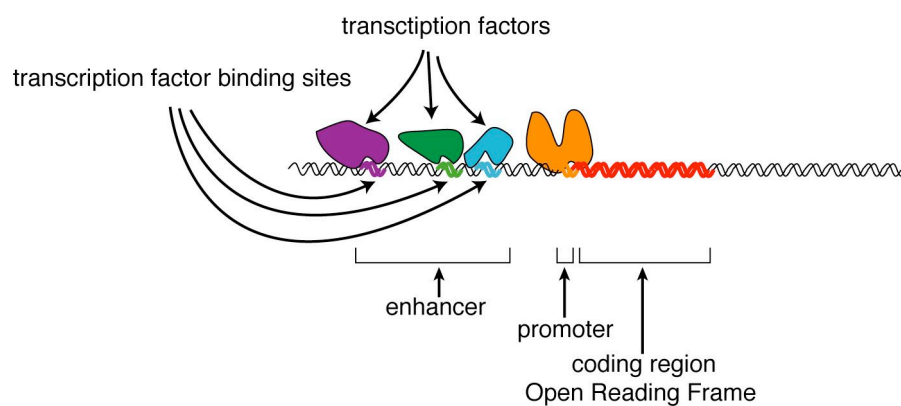


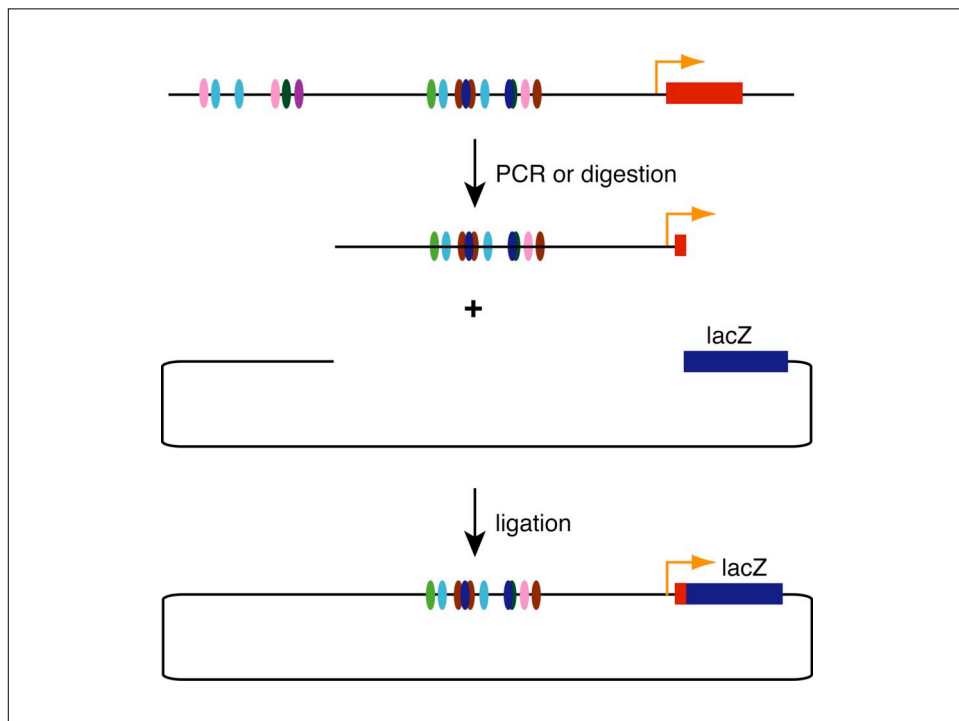
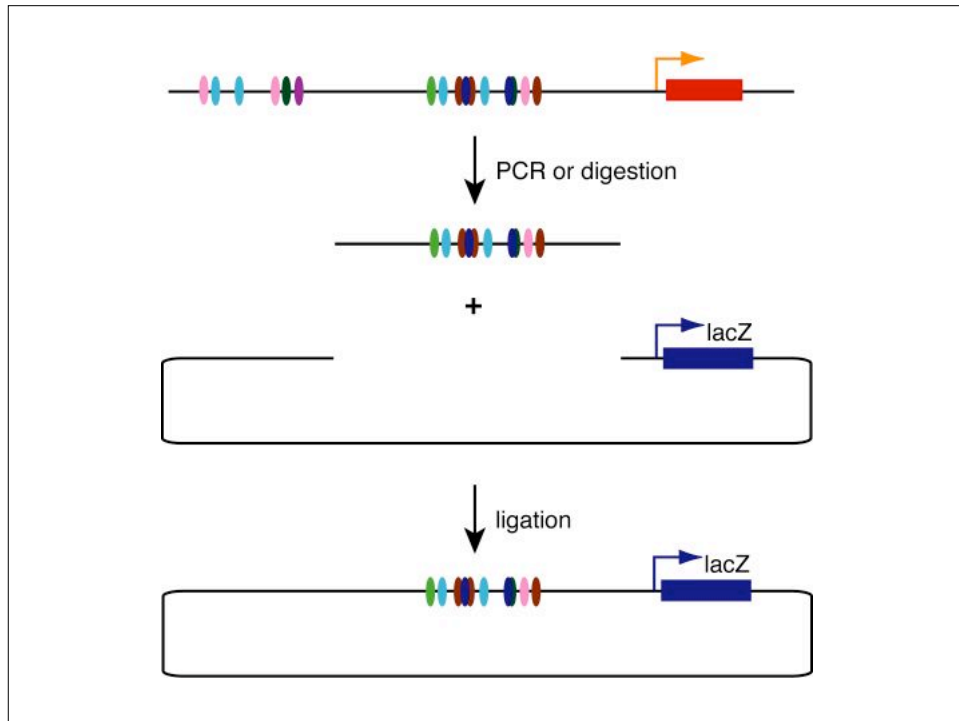
## Potential Issues

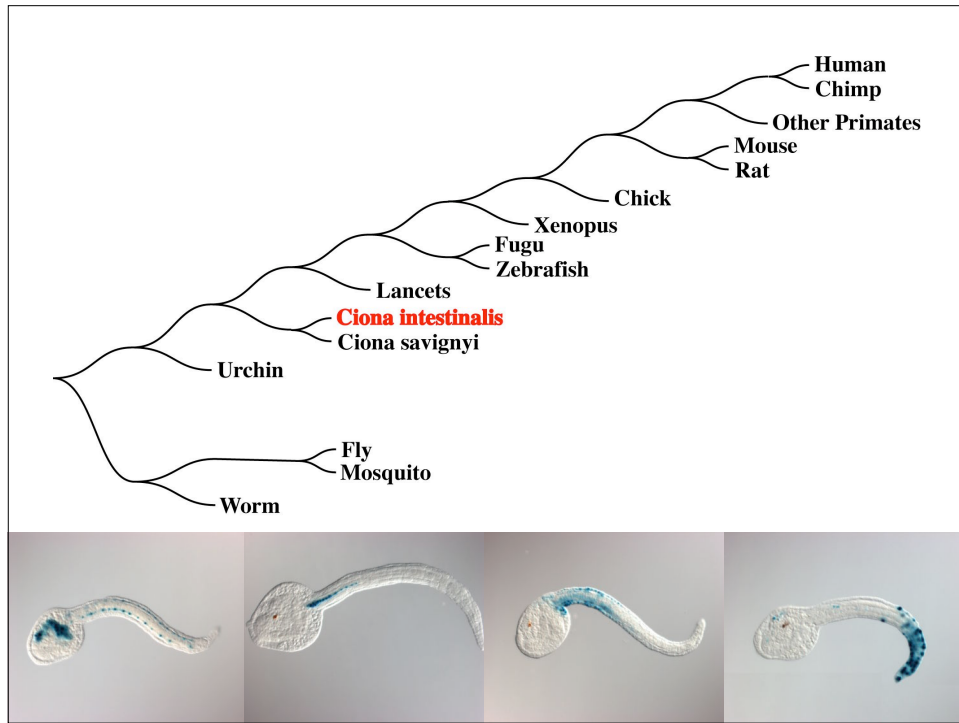
- Promoter specificity
- Insulators & repressors
- Enhancer Polarity
- Promoter competition
- Enhancers fragmented during cloning
  
- Timing
- Insufficient detection strength



## Characterize functional transcription factor binding sites







## A Functional Genomics Approach to Developmental Genetics

**Joint Genome Institute**  
**LBNL - LLNL - DOE**

**University of California**  
**Berkeley**

David Keys  
Byung-in Lee  
Chris Detter  
Stephan Trong

Sylvia Ahn  
Dave Engle  
Orsalem kahsai  
Maria Shin  
Joann Wang  
Mei Wang

Naoe Harafuji  
Anna Di Gregorio  
Mike Levine

